N° 82-003-X au catalogue
ISSN 1209-1375

Rapports sur la santé

Couplage des données du Recensement de 2006 et des données hospitalières au Canada

par Michelle Rotermann, Claudia Sanmartin, Richard Trudeau et Hélène St-Jean

Date de diffusion : le 21 octobre 2015

Statistique Canada Statistics Canada

Canada

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@canada.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros sans frais suivants :

- Service de renseignements statistiques 1-800-263-1136
- Service national d'appareils de télécommunications pour les malentendants 1-800-363-7629
- Télécopieur 1-877-287-4369

Programme des services de dépôt

- Service de renseignements 1-800-635-7943
- Télécopieur 1-800-565-7757

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « Normes de service à la clientèle ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, ses entreprises, ses administrations et les autres établissements. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Signes conventionnels dans les tableaux

Les signes conventionnels suivants sont employés dans les publications de Statistique Canada :

.	indisponible pour toute période de référence
..	indisponible pour une période de référence précise
...	n'ayant pas lieu de figurer
0	zéro absolu ou valeur arrondie à zéro
0^s	valeur arrondie à 0 (zéro) là où il y a une distinction importante entre le zéro absolu et la valeur arrondie
^{p}	provisoire
^{r}	révisé
x	confidentiel en vertu des dispositions de la *Loi sur la statistique*
^{E}	à utiliser avec prudence
F	trop peu fiable pour être publié
*	valeur significativement différente de l'estimation pour la catégorie de référence ($p<0,05$)

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2015

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'entente de licence ouverte de Statistique Canada.

Une version HTML **est aussi disponible.**

This publication is also available in English.

Couplage des données du Recensement de 2006 et des données hospitalières au Canada

par Michelle Rotermann, Claudia Sanmartin, Richard Trudeau et Hélène St-Jean

Résumé

Contexte : Le couplage d'enregistrements est couramment utilisé par les chercheurs du domaine de la santé pour combler les lacunes statistiques. L'étude présente un résumé des résultats du couplage des données du Recensement de la population de 2006 (sauf le Québec) et des données hospitalières provenant de la Base de données sur les congés des patients (BDCP).

Données et méthodes : On a utilisé l'appariement exact déterministe hiérarchique pour coupler les données du Recensement de 2006 et de la BDCP (2006-2007, 2007-2008 et 2008-2009), en fonction des clés de couplage dérivées de trois variables présentes dans les deux fichiers, à savoir la date de naissance, le code postal et le sexe. On a utilisé le fichier de données intégrales du recensement (questionnaire abrégé; 23,4 millions d'enregistrements) pour le couplage d'enregistrements; le fichier de données-échantillon (20 %) (questionnaire détaillé; 4,65 millions d'enregistrements), représentant la cohorte étudiée, a été utilisé pour la validation. Les fichiers couplés ont été comparés pour des secteurs de compétence, des années et d'autres covariables sélectionnées, pour ce qui est de l'admissibilité au couplage, des clés couplées et des taux de couplage et de couverture.

Résultats : Dans l'ensemble, 80 % des clés de couplage relevées dans la BDCP ont été couplées aux données du Recensement de 2006. Le pourcentage des répondants au recensement ayant rempli le questionnaire détaillé qui ont été couplés à au moins un dossier d'hospitalisation variait de 5 % à 8 % dans les différents secteurs de compétence; les taux de couplage étaient les plus élevés chez les grands utilisateurs connus de services hospitaliers, c'est-à-dire les groupes d'âge le plus avancé, les personnes à faible revenu et les Autochtones. En général, le fichier des données couplées englobe la majorité des événements ayant trait aux hospitalisations survenues au cours de la période visée par l'étude. Les taux de couverture pondérés et non pondérés variaient selon la région géographique et le groupe d'âge, les territoires et certains des groupes d'âge les plus jeunes présentant les taux pondérés les plus faibles.

Interprétation : Grâce à la méthode d'appariement exact déterministe hiérarchique, il est possible de coupler les données du recensement aux données relatives à plusieurs années dans la BDCP. L'incorporation des codes postaux mis à jour à partir des dossiers fiscaux a permis de réduire l'érosion du couplage au fil du temps. Les faibles taux de couverture observés pour les territoires et les groupes d'âge jeune laissent supposer que ces populations pourraient être sous-représentées dans les fichiers couplés.

Mots-clés : Couverture, collecte de données, couplage de données, bases de données, dossiers médicaux, dossiers d'hôpital.

Le couplage d'enregistrements, c'est-à-dire le processus d'appariement d'enregistrements entre différents ensembles de données ou à l'intérieur d'un seul ensemble, est couramment utilisé par les chercheurs dans le domaine de la santé pour combler les lacunes statistiques[1-7] et créer des ensembles de données ayant des applications diversifiées[6-11]. La majorité des couplages de données sur la santé au Canada se sont fondés sur les numéros d'assurance-maladie (NAM) extraits des registres provinciaux de l'assurance-maladie, qui sont uniques pour chaque personne[2,3,11,12]. Cependant, les NAM ne sont pas enregistrés dans la plupart des bases de données (p. ex., mortalité, recensement, impôt), et l'accès aux registres provinciaux est limité.

En l'absence d'un identificateur et/ou registre unique, une approche de rechange, soit l'appariement exact déterministe hiérarchique, peut servir au couplage des bases de données administratives sur la santé avec d'autres sources de données[13]. L'approche consiste à apparier différentes combinaisons d'identificateurs cooccurrents au niveau de la personne[14-18]. Statistique Canada a démontré la validité de cette approche en couplant les données du recensement et les données sur les hospitalisations pour deux provinces (l'Ontario et le Manitoba)[19]. En appariant les fichiers à partir de la date de naissance, du sexe et du code postal, on a obtenu des résultats similaires à ceux fondés sur les NAM extraits des registres provinciaux d'assurance-maladie[19].

La présente étude fait état des résultats d'un couplage par appariement exact hiérarchique de données du Recensement de la population de 2006 et de données hospitalières pour toutes les provinces et les territoires (sauf le Québec) provenant de la Base de données sur les congés des patients (BDCP) pour la période 2006-2007 à 2008-2009. Elle a pour but de déterminer si le couplage recensement—BDCP donne des résultats similaires d'un secteur de compétence à l'autre, et si les taux de couplage et de couverture diminuent en fonction du temps écoulé depuis le recensement. Le couplage a été approuvé par le Comité des politiques de Statistique Canada[20]. L'utilisation des données couplées est régie par la Directive sur le couplage d'enregistrements[21].

Auteurs : Michelle Rotermann (Michelle.Rotermann@statcan.gc.ca) et Claudia Sanmartin travaillent à la Division de l'analyse de la santé, Richard Trudeau (Richard.Trudeau@statcan.gc.ca), à la Division des enquêtes spéciales, et Hélène St-Jean, à la Division des méthodes d'enquêtes auprès des ménages de Statistique Canada, Ottawa (Ontario).

Sources des données

Recensement de la population de 2006

Les données du Recensement de 2006 ont été recueillies au moyen d'un questionnaire abrégé et d'un questionnaire détaillé ou complet. La plupart des ménages (80 %) ont reçu le questionnaire abrégé, qui comptait huit questions de base, y compris sur la date de naissance, le sexe et l'état matrimonial de tous les membres du ménage. Les autres ménages (20 %) ont reçu le questionnaire détaillé, qui comportait 53 autres questions portant sur des sujets comme la scolarité, l'ethnicité, la mobilité, le revenu et l'emploi[22]. Dans certaines régions, tous les ménages devaient remplir le questionnaire détaillé. C'était le cas au Nunavut, dans les Territoires du Nord Ouest (sauf Yellowknife), au Yukon (sauf Whitehorse) et dans les réserves indiennes et les établissements indiens[22]. Le recensement représente 95 % à 97 % de la population des provinces et 93 % à 94 % de la population des territoires[23].

Pour le couplage d'enregistrements, on a utilisé le fichier intégral du recensement (23,4 millions d'enregistrements), qui contient à la fois les enregistrements extraits du questionnaire abrégé et ceux tirés du questionnaire détaillé. La cohorte fondée sur les enregistrements pour le questionnaire détaillé (4,65 millions d'enregistrements) est celle utilisée dans la présente étude aux fins de validation.

Lorsque des inférences au sujet de la population canadienne sont faites à partir de données extraites du questionnaire détaillé, souvent les données sont pondérées[24]. Les poids d'échantillonnage permettent de tenir compte du plan de sondage ainsi que de la sous-représentation ou surreprésentation de personnes présentant certaines caractéristiques[25]. Les poids du recensement n'ont pas été ajustés en fonction de l'admissibilité pour le couplage.

La collecte et le traitement des données du recensement étaient soumis à des normes de qualité rigoureuses[22]. De plus, les données du recensement ont fait l'objet d'une vérification de la qualité, y compris de comparaisons avec d'autres sources de données. Les réponses incohérentes ou manquantes ont été imputées afin d'assurer la cohérence interne de l'information fournie par chaque ménage. Le taux d'imputation global a été de 2,9 % et les taux d'imputation des variables « âge » et « sexe » ont été inférieurs à 1,5 %.

Le succès de la collecte des données dépendait de l'exactitude des adresses, y compris les codes postaux. Dans les régions où les questionnaires du recensement avaient été distribués par la poste, Statistique Canada a validé et, au besoin, modifié les adresses avant le jour du recensement (16 mai 2006). Dans les régions où les questionnaires avaient été distribués par les agents recenseurs, les adresses figuraient dans une liste et ont été vérifiées au moment de la livraison.

Fichier maître sur les particuliers T1

Le Fichier maître sur les particuliers T1 (FMPT1) est un fichier annuel tiré des déclarations d'impôt. On y trouve le nom de la personne, la date de naissance, le sexe et le code postal, qui peuvent servir au couplage d'enregistrements. Pour les répondants au recensement dont le code postal était manquant, incomplet ou avait changé, on a couplé les FMPT1 pour 2005 à 2009 avec le fichier du Recensement de 2006 (questionnaire abrégé) de manière déterministe, à partir du sexe, de la date de naissance et des nom de famille et premier prénom partiels; les renseignements sur le revenu n'ont pas été utilisés. Environ 90 % des enregistrements du recensement ont été couplés à au moins un FMPT1 annuel. Dans le cas des personnes qui n'avaient pas produit une déclaration de revenus chaque année et(ou) qui n'étaient pas tenus d'en produire une (par exemple, les enfants), les codes postaux ont été attribués à partir de l'information sur les autres membres déclarants du ménage.

Base de données sur les congés des patients (BDCP)

La BDCP contient environ 3 millions de dossiers d'hôpital renfermant des données démographiques, administratives et cliniques, ainsi que les NAM, pour tous les congés reçus d'un établissement de soins de courte durée, et certains congés obtenus d'un hôpital psychiatrique, d'un hôpital pour malades chroniques, d'un établissement de réadaptation ou d'un service de chirurgies d'un jour au cours d'un exercice financier donné (1er avril au 31 mars) dans toutes les provinces, sauf le Québec, et les territoires[26,27]. Des études de seconde saisie des données, qui comparent certains champs des dossiers des patients avec les données correspondantes dans l'enregistrement original de la BDCP, ont trouvé à maintes reprises que les éléments de données non cliniques, y compris la date de naissance, le sexe et le code postal, sont d'une grande fiabilité, des différences ayant été observées dans ces champs dans moins de 2 % des cas[27].

Aux fins du couplage d'enregistrements, on a utilisé les données couvrant la période 2005-2006 à 2008-2009 de la BDCP pour la phase du prétraitement de la présente étude. Les données se rapportant à la période 2006-2007 à 2008-2009 ont été utilisées pour la phase du couplage des données (2006-2007 n = 3 186 079; 2007-2008 n = 3 204 838; 2008-2009 n = 3 232 396). Comme les enregistrements ont trait aux hospitalisations, non aux personnes, il est possible que les personnes hospitalisées à plusieurs reprises soient représentées plus d'une fois dans la BDCP.

Méthodes

Le couplage d'enregistrements comportait trois étapes, à savoir le traitement des données, le couplage d'enregistrements et la validation.

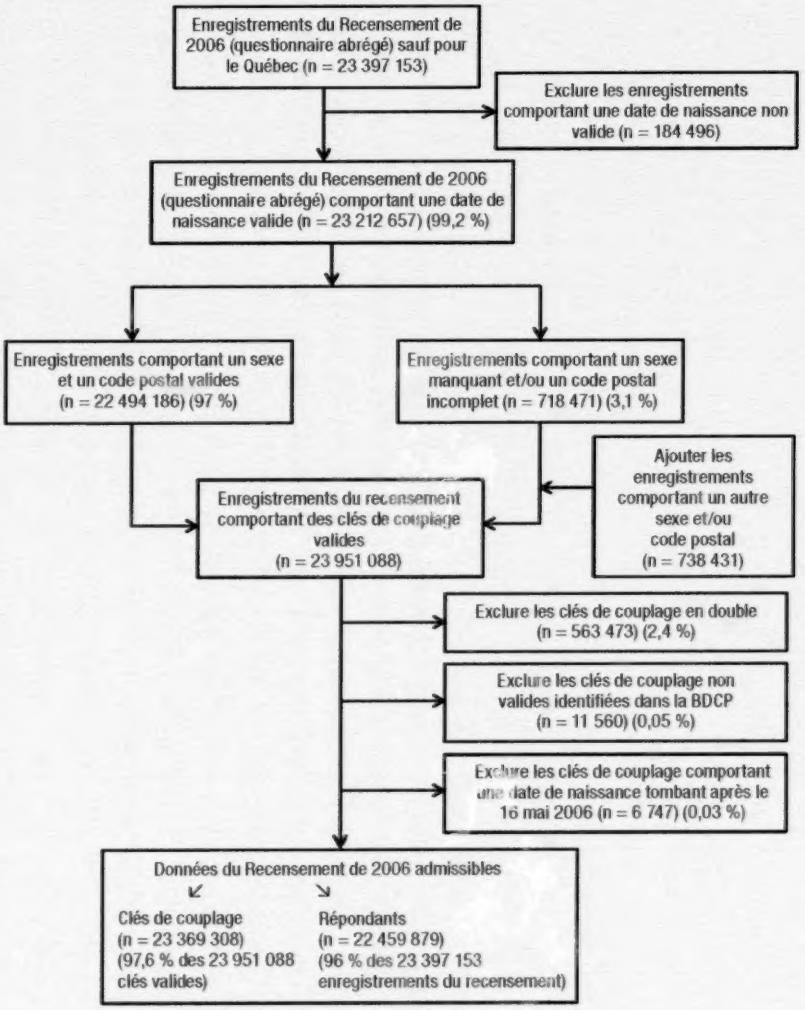
Traitement des données

Avant que ne soit effectué le couplage, les données ont été traitées afin d'améliorer la qualité des variables visées (date de naissance, code postal et sexe) et d'établir l'ensemble unique de clés de couplage dans chaque fichier de données, de manière à obtenir le moins de faux couplages possible. Ce traitement avait pour but de découvrir les erreurs ou les omissions dans les données qui pouvaient entraîner de faux couplages.

Au total, on disposait de 23 397 153 enregistrements du recensement pour le couplage[28] (figure 1). Les enregistrements pour le Québec ont été exclus parce que Statistique Canada n'a pas accès aux données de cette province sur les hospitalisations. Les enregistrements du recensement comportant une date de naissance non valide ou incomplète ont été exclus également. Lorsque le sexe manquait sur un enregistrement, on a attribué un sexe et on a créé un enregistrement en double comportant la même date de naissance et le même code postal, mais non le même sexe. Les enregistrements comportant un code postal différent dans le champ original et le champ post-traitement ont aussi été créés en double, le premier avec le code postal original, et l'autre avec le code post-traitement. Les enregistrements du recensement originaux et en double ont été associés au moyen d'un identificateur de groupe de recensement, ce qui a permis le repérage et le retrait des doublons après le couplage. Enfin, une série d'exclusions ont été appliquées aux enregistrements du recensement pour établir l'ensemble final de clés de couplage valides et uniques. Les clés suivantes ont été exclues : les clés en double comportant les mêmes date de naissance, code postal et sexe (par exemple, des jumeaux de même sexe habitant à la même adresse); les clés de couplage non valides repérées par suite du traitement des données de la BDCP (décrit ci-dessous); et les clés comportant une date de naissance tombant après le 16 mai 2006 (jour du recensement). Au total, 23 369 308 clés de recensement valides et uniques se prêtaient au couplage de données, ce qui représente 96 % des répondants au recensement (ceux du Québec non compris).

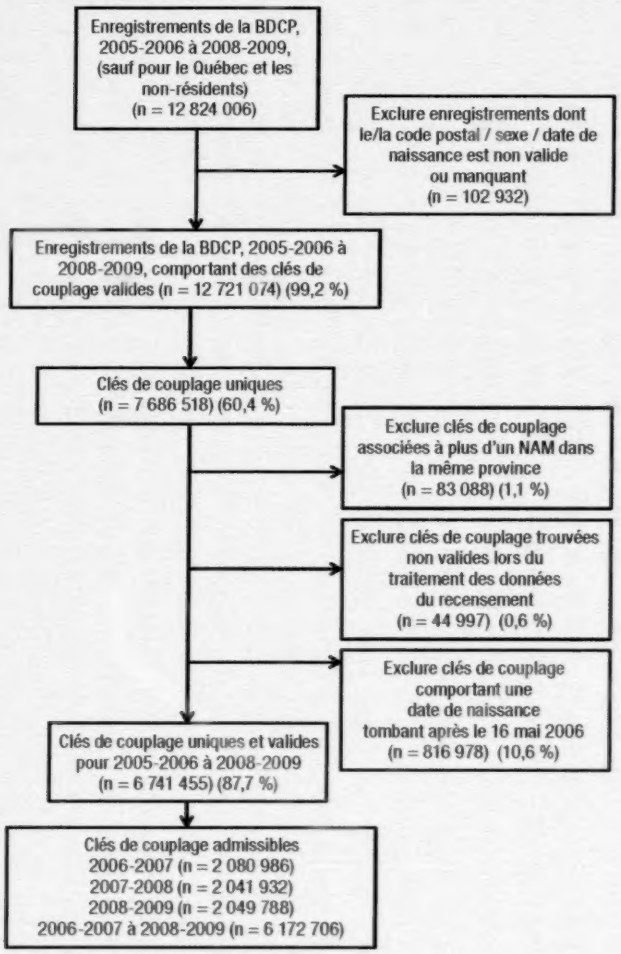
Les enregistrements de la BDCP ayant trait aux résidents canadiens hospitalisés dans d'autres provinces que le Québec et dans les territoires pour les exercices financiers 2005-2006 à 2008-2009 étaient admissibles pour un prétraitement (12 824 006) (figure 2). Les enregistrements comportant des données non valides ou manquantes pour la date de naissance, le sexe ou le code postal ont été exclus. Les enregistrements restants de la BDCP ont fait l'objet d'un traitement approfondi afin d'améliorer l'uniformité et l'exactitude des NAM et d'assurer une correspondance univoque avec les clés de couplage. Un ajustement couramment employé consiste à remplacer un NAM temporaire (normalement celui de la mère) ou absent pour un nouveau-né par un NAM permanent lui ayant été subséquemment attribué et ayant par la suite figuré dans les enregistrements d'hospitalisation.

Figure 1
Traitement des données du Recensement de 2006 préalable au couplage d'enregistrements



Plus de 12,7 millions d'enregistrements de la BDCP comportant des données valides aux fins de couplage ont été retenues, ce qui représente 7 686 518 clés de couplage uniques et valides (figure 2). Le grand nombre de clés de couplage en double n'avait rien de surprenant étant donné qu'une personne pouvait être hospitalisée à plusieurs reprises. Les clés suivantes ont été exclues : les clés de couplage associées à plus d'un NAM dans la même province; les clés non valides repérées par suite du

Figure 2
Traitement des données pour 2005-2006 à 2008-2009 de la Base de données sur les congés des patients en prévision du couplage d'enregistrements



traitement des données du recensement, décrit plus haut; et les clés comportant une date de naissance qui tombait après le jour du recensement. Au total, 6 172 706 clés de couplage uniques et valides se rapportant aux exercices 2006-2007 à 2008-2009 ont été retenues aux fins du couplage des données.

Couplage d'enregistrements

La méthode de couplage par appariement exact déterministe hiérarchique utilisée pour rapprocher les enregistrements du recensement et ceux de la BDCP comportait un processus itératif, selon lequel les clés de couplage indiquant une date de naissance, un sexe et un code postal étaient comparées entre fichiers. L'utilisation de clés multiples appliquées de façon consécutive permet de maximiser la puissance différenciatrice des données de couplage et de réduire au minimum l'incidence des erreurs et des données manquantes[13].

L'approche itérative intégrait 28 règles. Les premières itérations respectaient des règles rigoureuses, tandis que les subséquentes admettaient un certain degré de désaccord (tableau 1). Par exemple, la première itération exigeait une correspondance parfaite pour la date de naissance, le sexe et le code postal entre le fichier du recensement et les clés de couplage de la BDCP. Aux itérations 2 à 4, il devait y avoir une correspondance parfaite de la date de naissance, du sexe et du code postal entre le FMPT1 et les clés de couplage de la BDCP. On a utilisé les codes postaux dans les dossiers de l'impôt pour 2005, 2006 et 2007 lorsque le fichier du recensement s'appariait à la BDCP pour 2006-2007. Lorsqu'il s'appariait aux données de la BDCP pour 2007-2008 et pour 2008-2009, respectivement, on se référait respectivement aux dossiers de l'impôt pour 2006 à 2008 et pour 2007 à 2009. Aux itérations 5 à 10, les règles se rapportant au code postal étaient moins rigoureuses, de sorte que l'un des six caractères du code postal déclaré au recensement pouvait être omis. Ce processus a été repris en utilisant les codes postaux provenant des FMPT1 (itérations 11 à 28).

À chaque itération, seules les clés uniques de couplage d'un fichier étaient comparées aux clés uniques de l'autre fichier. Une fois une itération terminée, les clés couplées étaient soustraites des itérations ultérieures de manière que chaque clé ne soit couplée qu'une fois. Les clés couplées qui avaient été ajoutées pour tenir compte de données manquantes pour le sexe et(ou) le code postal et qui avaient le même identificateur de groupe de recensement ont également été retirées à cette étape. Lorsqu'une clé du recensement était couplée à une clé de la BDCP, on appliquait un couplage déterministe à l'ensemble des enregistrements de congé d'hôpital de la BDCP, à partir de la clé de couplage, du NAM correspondant et de la province de délivrance. Le traitement des données et le couplage d'enregistrements ont été exécutés en SAS, v. 9.2.

Validation

Deux types de taux de couplage sont présentés. Premièrement, pour chaque itération du couplage, on présente le pourcentage des clés de la BDCP qui ont été couplées avec une clé du fichier de recensement. On ne s'attendait pas à

ce que le taux de couplage final soit de 100 %, compte tenu des différences entre les populations représentées dans le fichier du recensement et le fichier de la BDCP, du sous-dénombrement au recensement de certaines sous-populations et de l'utilisation des services hospitaliers par les personnes entrées au Canada après le jour du recensement (p. ex. les nouveaux immigrants). On s'attendait à observer des taux de couplage faibles également chez les personnes qui vivaient en établissement (p. ex., un établissement de soins de longue durée), puisque ces personnes se trouveraient à partager le même code postal, ce qui pourrait réduire la qualité des résultats ayant trait à l'unicité des clés de couplage.

Deuxièmement, le pourcentage des enregistrements du recensement qui étaient admissibles pour le couplage qui ont été couplés à la BDCP (2006-2007 à 2008-2009) sont présentés selon la province ou le territoire, pour certaines caractéristiques sociodémographiques. Ces taux sont fondés sur les personnes ayant rempli le questionnaire détaillé du recensement (la cohorte de validation) qui étaient admissibles pour le couplage (4 652 683 enregistrements, ce qui représente 94 % des questionnaires détaillés) et reflètent la prévalence d'au moins une hospitalisation au cours de l'exercice financier respectif (tableau 3). On s'attendait à ce que les taux de couplage soient plus élevés dans certains groupes de population, par exemple les aînés, qui sont plus susceptibles d'être hospitalisés en comparaison des personnes non âgées. La variation observée dans les taux de couplage d'une province ou d'un territoire à l'autre pourrait aussi refléter des différences relativement à la prestation des soins de santé[29].

Analyse de la couverture

Lorsqu'une fraction seulement des enregistrements seront présumément couplés, l'évaluation de la qualité d'un couplage de données représente un défi, ce qui arrive typiquement dans le cas de couplages de données sur la santé où, par exemple, le nombre attendu de personnes qui seront hospitalisées ou décéderont au cours de la période de suivi est limité. Dans de telles situations, il n'est pas facile de savoir si un enregistrement non couplé représente un couplage manqué ou si l'événement d'intérêt ne s'est pas produit. Pour évaluer la qualité de couplages de ce genre, on a comparé les résultats d'approches différentes fondées sur les mêmes données[16,17,30-33] et comparé les taux de résultat et les répartitions en pourcentage pour des variables disponibles dans des données couplées et des données non couplées[30,34,35].

On a calculé les taux de couverture annuels non pondérés et pondérés à l'échelle nationale (le Québec non compris) et selon le secteur de compétence, en divisant le nombre de sorties d'hôpital (établissements de soins de courte durée) chez les personnes ayant rempli le questionnaire détaillé du recensement dans chaque secteur de compétence, d'après les données tirées du couplage recensement–BDCP (numérateur), par le nombre de sorties d'hôpital (établissements de soins de courte durée) d'après les données non couplées de la BDCP pour 2006-2007, 2007-2008 et 2008-2009 (dénominateur). Afin d'assurer une concordance plus étroite avec les populations visées par les données couplées, dans la mesure du possible, on a retiré du dénominateur les enregistrements de la BDCP se rapportant aux populations non saisies à l'aide du questionnaire détaillé du recensement, c'est-à-dire les résidents de foyers pour personnes âgées, les personnes nées après le jour du recensement, les morts-nés et les non-Canadiens.

Les taux de couverture non pondérés devraient se rapprocher du pourcentage de la population ayant rempli le questionnaire détaillé (environ 20 % à l'échelle nationale, variant de 16 % pour Terre-Neuve-et-Labrador à 63 % à 69 % pour le Nunavut)[22,24]. Les taux de couverture pondérés devraient être voisins de 100 % sans toutefois l'atteindre, compte tenu des différences entre les populations représentées par les données couplées et celles reflétées dans les données non couplées de la BDCP. Par exemple, la population vivant en établissement, un grand consommateur de services hospitaliers[36,37], est prise en compte dans les données non couplées de la BDCP, mais non dans les données tirées du couplage recensement–BDCP.

Caractéristiques sociodémographiques

Les taux de couplage ont été comparés selon certaines caractéristiques socioéconomiques visées par le recensement afin de déterminer s'ils étaient plus élevés chez les personnes susceptibles d'être hospitalisées, comme les personnes à faible revenu et les Autochtones[36,38,39].

Les quintiles de revenu ont été calculés au niveau de la famille économique ou établis directement pour les personnes seules[23]. On a calculé le revenu total après impôt en additionnant le revenu provenant de toutes les sources pour tous les membres de la famille ou pour la personne seule. Celui-ci a ensuite été ajusté en fonction de la taille de la famille, puis réparti en fonction de *quintiles*. Afin de réduire au minimum les différences régionales de revenu, on a estimé les quintiles pour chaque province ou territoire séparément, puis combiné les résultats.

Le *niveau de scolarité le plus élevé* dans le cas des personnes de 18 ans et plus a été réparti en fonction de deux catégories, à savoir diplôme d'études secondaires ou plus, et pas de diplôme d'études secondaires. Les personnes de moins de 18 ans, pour la plupart trop jeunes pour détenir un diplôme, ont été exclues.

L'information relative au *statut d'Autochtone* se fonde sur la question « Cette personne est-elle un Autochtone, c'est-à-dire un Indien de l'Amérique du Nord, un Métis ou un Inuit (Eskimau)? ». Les répondants au recensement devaient cocher toutes les réponses qui s'appliquaient. Les réponses ont été regroupées en six catégories, à savoir Indien de l'Amérique du Nord (seulement), Métis (seulement), Inuit (seulement), autre Autochtone (origines multiples ou indéterminées), Autochtone (combinaison d'origines parmi les

Tableau 1
Pourcentage et nombre cumulatifs de clés de la Base de données sur les congés des patients (BDCP) couplées aux données du Recensement de 2006, par itération, Canada (sauf le Québec), 2006-2007, 2007-2008 et 2008-2009

	BDCP pour 2006-2007			BDCP pour 2007-2008			BDCP pour 2008-2009		
	Source du code postal (clé du recensement seulement) et format si incomplet	Cumulatif		Source du code postal (clé du recensement seulement) et format si incomplet	Cumulatif		Source du code postal (clé du recensement seulement) et format si incomplet	Cumulatif	
Code postal utilisé / Itération		%	nombre		%	nombre		%	nombre
Codes postaux complets tirés du recensement (RTA UDL)									
1	recensement	71,8	1 495 004	recensement	65,1	1 329 408	recensement	59,4	1 217 899
Codes postaux complets tirés du FMPT1 (RTA UDL)									
2	FMPT1 2005	74,3	1 545 888	FMPT1 2006	72,0	1 470 679	FMPT1 2007	70,3	1 440 860
3	FMPT1 2006	75,5	1 570 658	FMPT1 2007	73,8	1 506 694	FMPT1 2008	72,0	1 475 942
4	FMPT1 2007	75,7	1 574 321	FMPT1 2008	74,1	1 512 212	FMPT1 2009	72,3	1 482 647
Codes postaux partiels tirés du recensement (un x indique une valeur alphanumérique omise de la clé)									
5	RTA UDx	76,5	1 591 750	RTA UDx	74,8	1 527 771	RTA UDx	73,0	1 496 968
6	RTA UxL	77,9	1 621 704	RTA UxL	76,3	1 557 111	RTA UxL	74,4	1 525 707
7	RTA xDL	78,2	1 627 811	RTA xDL	76,5	1 562 431	RTA xDL	74,7	1 530 638
8	RTx UDL	78,9	1 642 783	RTx UDL	77,2	1 575 974	RTx UDL	75,3	1 543 716
9	RxA UDL	79,0	1 644 879	RxA UDL	77,3	1 577 857	RxA UDL	75,4	1 545 438
10	xTA UDL	79,2	1 648 534	xTA UDL	77,5	1 581 599	xTA UDL	75,6	1 549 564
Codes postaux partiels tirés du FMPT1 (2005, 2006, 2007) pour l'année de la BDCP - 1									
11	RTA UDx	79,3	1 649 501	RTA UDx	77,6	1 583 960	RTA UDx	75,8	1 553 001
12	RTA UxL	79,4	1 651 674	RTA UxL	77,8	1 587 924	RTA UxL	76,0	1 558 490
13	RTA xDL	79,4	1 652 018	RTA xDL	77,8	1 588 708	RTA xDL	76,1	1 559 725
14	RTx UDL	79,4	1 652 996	RTx UDL	77,9	1 590 642	RTx UDL	76,2	1 562 475
15	RxA UDL	79,4	1 653 106	RxA UDL	77,9	1 590 959	RxA UDL	76,2	1 562 937
16	xTA UDL	79,5	1 653 459	xTA UDL	77,9	1 591 447	xTA UDL	76,3	1 563 595
Codes postaux partiels tirés du FMPT1 (2006, 2007, 2008) pour l'année de la BDCP									
17	RTA UDx	79,5	1 654 162	RTA UDx	78,0	1 592 392	RTA UDx	76,3	1 564 602
18	RTA UxL	79,6	1 655 504	RTA UxL	78,1	1 593 978	RTA UxL	76,4	1 566 193
19	RTA xDL	79,6	1 655 723	RTA xDL	78,1	1 594 306	RTA xDL	76,4	1 566 493
20	RTx UDL	79,6	1 656 310	RTx UDL	78,1	1 595 077	RTx UDL	76,5	1 567 262
21	RxA UDL	79,6	1 656 404	RxA UDL	78,1	1 595 219	RxA UDL	76,5	1 567 402
22	xTA UDL	79,6	1 656 620	xTA UDL	78,1	1 595 467	xTA UDL	76,5	1 567 654
Codes postaux partiels tirés du FMPT1 (2007, 2008, 2009) pour l'année de la BDCP + 1									
23	RTA UDx	79,6	1 657 049	RTA UDx	78,2	1 595 918	RTA UDx	76,5	1 568 170
24	RTA UxL	79,7	1 657 975	RTA UxL	78,2	1 596 983	RTA UxL	76,6	1 569 308
25	RTA xDL	79,7	1 658 107	RTA xDL	78,2	1 597 128	RTA xDL	76,6	1 569 469
26	RTx UDL	79,7	1 658 563	RTx UDL	78,2	1 597 581	RTx UDL	76,6	1 569 988
27	RxA UDL	79,7	1 658 619	RxA UDL	78,2	1 597 653	RxA UDL	76,6	1 570 075
28	xTA UDL	79,7	1 658 817	xTA UDL	78,3	1 597 868	xTA UDL	76,6	1 570 329

Notes : Les codes postaux comptent six caractères. Les trois premiers (RTA) se rapportent à la région de tri d'acheminement et les trois derniers (UDL), à l'unité de distribution locale. Un « x » à la place d'un caractère du code postal (RTA UDL) indique que ce caractère a été omis dans la clé de couplage. Chaque clé renferme une date de naissance, un sexe et un code postal. Les clés de la BDCP tenaient compte du code postal de la résidence du patient et étaient uniformes pour les 28 itérations. Les clés du recensement renfermaient les codes postaux déclarés au recensement et d'autres tirés des fichiers d'impôt (FMPT1). La source et l'intégralité du code postal variaient avec chaque itération. Le nombre de clés de la BDCP qui étaient admissibles pour le couplage était de 2 080 986 (2006-2007), de 2 041 932 (2007-2008) et de 2 049 788 (2008-2009).

quatre catégories précédentes), et non-Autochtone.

Le pays de naissance, la citoyenneté et le statut d'immigrant ont été combinés en une variable, soit *statut d'immigrant* (immigrant, non-immigrant ou résident non permanent). Les immigrants ont ensuite été classés en immigrants *de longue date* (arrivés au moins 10 ans avant le Recensement de 2006) et immigrants *récents* (arrivés tout au plus neuf ans avant le Recensement de 2006).

Une variable sur la *mobilité* (lieu de résidence un an auparavant) a été créée pour rendre compte des changements

d'adresse (même adresse, autre adresse au Canada ou autre adresse à l'extérieur du Canada). La variable a été dérivée en comparant la municipalité et la province de résidence du recensé le jour du recensement et un an plus tôt.

Une variable sur le *milieu rural/urbain* reflétait l'emplacement et la taille de la collectivité. Étaient considérées habiter en milieu rural/agricole les personnes vivant dans une collectivité de moins de 1 000 habitants, que ce soit dans une ferme ou non. Les centres de population pouvaient être petits (1 000 à 29 999 habitants), moyens (30 000 à 99 999 habitants) ou grands (100 000 habitants ou plus).

Protection de la confidentialité des données du répondant

Statistique Canada assure la protection de la confidentialité des données du répondant pendant et après le couplage. Seuls les employés que le couplage concerne directement ont accès aux identificateurs uniques requis pour le couplage (p. ex. le nom et le numéro d'assurance-maladie) et ils *ne* voient *pas* les données sur la santé. Une fois le couplage de données terminé, on crée un fichier analytique duquel on soustrait les identificateurs personnels. Ce fichier sans identificateurs est utilisé par les analystes pour la validation et l'analyse.

Tableau 2
Pourcentage et nombre de clés de la Base de données sur les congés des patients (BDCP) couplées aux données du Recensement de 2006 et répartition en pourcentage des clés couplées, selon l'itération, par province, sexe et groupe d'âge, Canada (sauf le Québec), 2006-2007, 2007-2008 et 2008-2009

	2006-2007					2007-2008					2008-2009				
			Pourcentage de clés couplées, par itération					Pourcentage de clés couplées, par itération					Pourcentage de clés couplées, par itération		
	%	Nombre	1	2 à 4	5 à 28	%	Nombre	1	2 à 4	5 à 28	%	Nombre	1	2 à 4	5 à 28
Canada (sauf le Québec)	79,7	1 658 817	90,1	4,8	5,1	78,3	1 597 868	83,2	11,4	5,4	76,6	1 570 329	77,6	16,9	5,6
Province/Territoire															
Terre-Neuve-et-Labrador	84,5	66 832	91,6	3,9	4,6	84,7	66 880	87,9	7,8	4,3	83,8	66 186	83,8	11,6	4,6
Île-du-Prince-Édouard	83,0	16 162	90,0	4,5	5,5	82,4	15 553	84,9	9,1	6,0	82,3	15 667	81,4	13,2	5,4
Nouvelle-Écosse	84,0	99 818	90,9	4,7	4,4	83,3	99 493	86,0	9,5	4,6	82,1	100 570	81,1	14,0	4,9
Nouveau-Brunswick	80,8	78 694	91,7	5,2	3,1	79,5	74 791	85,5	11,1	3,4	78,1	72 768	80,7	16,2	3,1
Ontario	80,3	577 266	91,1	4,2	4,7	78,6	545 688	83,7	11,3	5,0	76,5	527 613	77,7	16,9	5,3
Manitoba	81,2	119 511	90,3	4,3	5,5	79,9	114 433	84,9	9,6	5,5	79,0	114 755	79,8	14,4	5,8
Saskatchewan	81,6	121 636	89,7	5,2	5,1	80,6	118 397	83,9	10,9	5,2	79,6	118 056	78,5	15,9	5,6
Alberta	77,2	177 466	87,9	6,3	5,7	75,1	168 781	79,0	14,8	6,2	72,6	162 203	72,2	21,2	6,6
Colombie-Britannique	78,0	391 169	89,1	5,0	5,8	76,5	383 933	81,7	12,2	6,1	75,0	382 745	76,0	17,9	6,1
Yukon	73,8	2 817	85,6	9,2	5,1	71,6	2 754	78,5	15,2	6,3	69,7	2 742	72,1	21,5	6,4
Territoires du Nord-Ouest / Nunavut	73,1	7 242	84,1	6,0	10,0	73,4	6 887	79,8	10,0	10,2	73,0	6 672	76,9	12,4	10,7
Sexe															
Hommes	79,0	675 424	90,5	4,2	5,4	78,5	652 339	85,2	9,3	5,5	77,6	646 688	80,8	13,7	5,5
Femmes	80,2	983 393	89,9	5,2	4,9	78,1	945 529	81,8	12,9	5,3	75,9	923 641	75,3	19,1	5,6
Groupe d'âge															
Moins de 1 an	73,3	37 835	91,3	3,8	4,9	69,8	9 237	72,9	19,5	7,6	67,5	7 313	63,1	27,9	9,1
1 an à 4 ans	78,7	35 894	87,3	6,0	6,7	76,2	29 564	78,0	15,1	6,9	74,1	25 841	70,6	22,2	7,2
5 à 9 ans	81,9	27 034	89,2	4,7	6,2	78,6	22 638	81,9	11,6	6,5	77,1	20 166	75,5	17,5	7,0
10 à 14 ans	81,6	25 870	89,3	4,4	6,2	78,9	28 664	82,6	10,3	7,0	76,1	32 349	77,7	15,2	7,1
15 à 19 ans	72,8	49 831	85,1	6,5	8,4	68,1	52 379	74,5	15,9	9,6	65,0	54 298	66,5	23,2	10,4
20 à 24 ans	67,6	70 637	79,8	11,2	8,9	62,2	70 224	63,1	26,6	10,3	59,0	70 808	51,2	37,6	11,2
25 à 34 ans	77,9	225 732	86,4	8,1	5,5	73,4	217 485	71,6	22,2	6,2	69,9	205 788	60,0	33,1	6,9
35 à 44 ans	80,1	191 606	89,0	5,2	5,7	77,7	180 773	80,9	12,9	6,2	76,0	173 808	74,5	19,1	6,4
45 à 54 ans	82,5	214 923	90,7	4,0	5,3	81,7	218 754	85,6	9,0	5,4	81,0	224 727	81,2	13,3	5,5
55 à 64 ans	84,6	239 176	92,0	3,6	4,4	84,5	245 326	87,9	7,8	4,3	83,9	252 859	84,1	11,7	4,2
65 à 74 ans	84,0	234 975	93,2	3,1	3,7	84,3	237 871	89,9	6,5	3,6	83,8	240 058	86,8	9,6	3,6
75 ans et plus	77,5	305 304	92,9	3,1	4,0	78,0	284 953	89,5	6,3	4,1	76,4	262 314	86,2	9,7	4,1

Notes : Le sexe et le groupe d'âge ont été définis à partir de données dans la BDCP et dans la clé de couplage seulement (composée des variables date de naissance / sexe / code postal). La province/territoire a été définie d'après la première lettre du code postal. Les codes postaux pour les Territoires du Nord-Ouest et le Nunavut commençant tous par la lettre X, les données pour ces territoires ont été combinées.
Sources : Recensement de la population de 2006; fichier de données résultant du couplage recensement–BDCP pour 2006-2007 à 2008-2009.

Résultats

Résultats du couplage

En tout, 80 % (1,66 million) des clés de la BDCP admissibles pour 2006-2007 ont été couplées aux données du Recensement de 2006 (tableau 1). Les résultats étaient similaires pour l'exercice 2007-2008 de la BDCP (78 % des clés admissibles, ou 1,60 million) et pour 2008-2009 (77 % des clés admissibles, ou 1,57 million). La majorité des couplages (72 % pour 2006-2007 à 59 % pour 2008-2009, ou 1,50 à 1,22 million, respectivement) ont été obtenus à la première itération, pour laquelle il devait y avoir une correspondance parfaite de la date de naissance, du sexe et du code postal. Le nombre de couplages obtenus à partir des codes postaux dans les dossiers d'impôt (itérations 2 à 4) variait de 79 000 (4 %) pour 2006-2007 à 265 000 (13 %) pour 2007-2008. Les itérations 5 à 28 ont donné lieu à 85 000 à 88 000 (4 %) autres couplages, selon l'année de la BDCP.

Les pourcentages de clés de la BDCP qui ont été couplées avaient tendance à être uniformes d'une province à l'autre et d'un groupe d'âge à l'autre, à quelques exceptions près (tableau 2). Par exemple, les pourcentages étaient comparativement faibles en Alberta (77 % pour l'exercice 2006-2007 de la BDCP à 73 % pour 2008-2009), en Colombie-Britannique (78 % à 75 %) et dans les territoires (73 % à 72 %). On a également observé des taux faibles chez les nourrissons de moins de 1 an (73 % à 68 %) et chez les 15 à 24 ans (70 % à 61 %).

Le pourcentage des répondants au recensement dénombrés à l'aide du questionnaire détaillé qui ont fait l'objet d'un couplage avec la BDCP (c'est-à-dire, qui avaient été hospitalisés) variait de 5,6 % (2006-2007) à 5,2 % (2008-2009) (tableau 3). Les taux de couplage reflétaient les écarts attendus d'utilisation des services hospitaliers. Le taux chez les femmes était plus élevé que chez les hommes. Les nourrissons qui avaient moins de 1 an le jour du recensement

Tableau 3
Nombre et pourcentage de répondants au Recensement de 2006 (questionnaire détaillé) admissibles pour le couplage et ayant obtenu leur congé d'un hôpital de soins de courte durée, selon certaines caractéristiques, Canada (sauf le Québec), 2006-2007, 2007-2008 et 2008-2009

	Répondants (questionnaire détaillé)	Répondants admissibles pour le couplage		Répondants au recensement admissibles couplés à la BDCP					
				2006-2007		2007-2008		2008-2009	
	milliers	milliers	%	milliers	%	milliers	%	milliers	%
Canada (sauf le Québec)	4 946,2	4 652,7	94	259,7	5,6	249,1	5,4	241,5	5,2
Province/Territoire									
Terre-Neuve-et-Labrador	103,5	97,8	94	5,9	6,1	5,9	6,0	5,7	5,8
Île-du-Prince-Édouard	26,2	24,5	94	1,8	7,2	1,7	6,9	1,6	6,6
Nouvelle-Écosse	182,4	171,6	94	9,7	5,6	9,4	5,5	9,2	5,4
Nouveau-Brunswick	146,5	141,2	96	10,0	7,1	9,5	6,8	9,0	6,4
Ontario	2 387,9	2 254,5	94	114,2	5,1	109,3	4,8	105,9	4,7
Manitoba	278,9	263,7	94	17,4	6,6	16,5	6,3	16,6	6,3
Saskatchewan	239,2	225,2	94	17,6	7,8	16,5	7,3	15,9	7,1
Alberta	681,5	640,7	94	37,3	5,8	35,8	5,6	34,4	5,4
Colombie-Britannique	830,5	769,1	92	41,6	5,4	40,4	5,3	39,2	5,1
Yukon	13,9	13,0	94	0,7	5,3	0,7	5,5	0,7	5,7
Territoires du Nord-Ouest	26,3	24,6	93	1,7	6,8	1,6	6,5	1,5	6,2
Nunavut	29,3	26,8	91	1,7	6,5	1,7	6,4	1,7	6,4
Sexe									
Hommes	2 432,5	2 284,4	94	101,8	4,5	97,5	4,3	95,5	4,2
Femmes	2 513,7	2 368,3	94	157,9	6,7	151,5	6,4	146,0	6,2
Groupe d'âge									
Moins de 1 an	56,8	51,8	91	8,4	16,2	1,7	3,4	1,1	2,1
1 an à 4 ans	229,8	213,0	93	5,9	2,7	4,2	2,0	3,4	1,6
5 à 9 ans	305,2	284,4	93	4,2	1,5	3,7	1,3	3,4	1,2
10 à 14 ans	345,1	321,7	93	4,7	1,5	5,5	1,7	6,1	1,9
15 à 19 ans	354,7	330,7	93	9,8	2,9	10,4	3,2	10,9	3,3
20 à 24 ans	329,6	307,7	93	13,8	4,5	14,1	4,6	13,7	4,5
25 à 34 ans	623,4	587,4	94	41,7	7,1	40,2	6,8	37,2	6,3
35 à 44 ans	756,2	715,5	95	29,1	4,1	26,7	3,7	24,9	3,5
45 à 54 ans	767,2	725,0	94	28,8	4,0	28,9	4,0	29,3	4,0
55 à 64 ans	556,1	525,3	94	33,0	6,3	34,0	6,5	34,1	6,5
65 à 74 ans	343,6	325,9	95	34,5	10,6	35,2	10,8	35,3	10,8
75 ans et plus	278,5	264,2	95	46,0	17,4	44,4	16,8	42,0	15,9
Quintile de revenu du ménage									
1 (le plus faible)	989,2	911,6	92	56,7	6,2	53,0	5,8	50,6	5,6
2	989,2	927,1	94	58,7	6,3	56,5	6,1	54,6	5,9
3	989,3	933,6	94	51,5	5,5	49,8	5,3	48,5	5,2
4	989,3	938,1	95	47,8	5,1	46,2	4,9	44,9	4,8
5 (le plus élevé)	989,3	942,1	95	44,9	4,8	43,6	4,6	42,9	4,6
Niveau de scolarité (18 ans et plus)									
Diplôme d'études secondaires ou plus	2 979,1	2 818,5	95	160,0	5,7	158,1	5,6	153,6	5,5
Pas de diplôme d'études secondaires	811,9	759,9	93	71,2	9,4	69,8	9,2	67,5	8,9
Identité autochtone									
Toute identité autochtone	470,2	422,2	90	32,6	7,7	31,0	7,3	30,2	7,2
Indien de l'Amérique du Nord seulement	345,0	306,4	89	25,0	8,1	23,8	7,8	23,3	7,6
Métis seulement	84,3	78,4	93	5,0	6,4	4,6	5,9	4,5	5,7
Inuit seulement	32,5	29,9	92	2,1	7,0	2,0	6,9	2,0	6,7
Identités autochtones multiples	8,4	7,6	91	0,5	6,8	0,5	6,9	0,5	6,6
Non-Autochtone	4 475,9	4 230,4	94	227,1	5,4	218,1	5,2	211,3	5,0

(16 mai 2006) et les personnes âgées constituaient les groupes d'âge où l'on était le plus susceptible de faire l'objet d'un couplage avec la BDCP pour l'exercice 2006-2007. Au cours des exercices de la BDCP subséquents, le taux de couplage chez les personnes âgées est demeuré comparativement élevé, mais chez les moins de 1 an en date du jour du recensement, il a reculé, rejoignant celui observé pour les enfants qui étaient âgés de 1 à 4 ans le jour du recensement. Comptaient parmi d'autres groupes ayant des taux de couplage élevés les personnes appartenant à un quintile de faible revenu (6 %) et les personnes ayant déclaré être d'identité autochtone (7 %). En général, les taux de couplage avaient tendance à être plus élevés en milieu rural qu'en milieu urbain.

Évaluation de la couverture

Les taux de couverture associés aux hospitalisations toutes causes confondues pour 2006-2007 à 2008-2009 ont été de 17 % (sans pondération) et de 80 % à 78 % (avec pondération) (tableau 4), mais ils variaient selon le secteur de compétence. Par exemple, les taux de couverture non pondérés variaient de 16 % à 22 % dans les provinces, et de 29 % à 69 % dans les territoires. Les taux de couverture pondérés variaient de 75 % à 84 % dans les provinces, et de 62 % à 72 % dans les territoires.

Au cours de la période visée par l'étude, les taux de couverture non pondérés et pondérés étaient similaires par sexe, mais non par groupe d'âge. Les taux pondérés pour 2006-2007 pour les moins de 1 an et pour les 15 à 24 ans étaient de 6 à 10 points de pourcentage inférieurs au total tous âges confondus. Les taux de couverture pondérés fondés sur les fichiers couplés pour 2007-2008 et pour 2008-2009 reflétaient cette tendance également, mais comme l'âge était défini en fonction du recensement plutôt que d'après les données sur les hospitalisations, le sous-dénombrement des jeunes dans l'avant-dernier et le dernier fichier de la BDCP s'observait chez les groupes d'âge qui suivaient.

Tableau 3 (fin)
Nombre et pourcentage de répondants au Recensement de 2006 (questionnaire détaillé) admissibles pour le couplage et ayant obtenu leur congé d'un hôpital de soins de courte durée, selon certaines caractéristiques, Canada (sauf le Québec), 2006-2007, 2007-2008 et 2008-2009

	Répondants (questionnaire détaillé)	Répondants admissibles pour le couplage		Répondants au recensement admissibles couplés à la BDCP					
				2006-2007		2007-2008		2008-2009	
	milliers	milliers	%	milliers	%	milliers	%	milliers	%
Statut d'immigrant									
Non-immigrant	3 864,2	3 627,2	94	205,5	5,7	196,2	5,4	190,3	5,2
Immigrant	1 034,8	982,5	95	53,1	5,4	51,9	5,3	50,2	5,1
De longue date (arrivé au moins 10 ans avant le Recensement de 2006)	315,5	297,9	94	11,9	4,0	11,1	3,7	10,3	3,4
Récent (arrivé tout au plus 9 ans avant le Recensement de 2006)	719,3	684,6	95	41,1	6,0	40,8	6,0	40,0	5,8
Résident non permanent	47,0	42,9	91	1,1	2,5	1,0	2,4	1,0	2,3
Mobilité résidentielle									
Même adresse	4 181,3	3 943,4	94	217,5	5,5	214,5	5,4	208,6	5,3
Autre adresse au Canada	651,9	610,5	94	32,2	5,3	31,3	5,1	30,3	5,0
Autre adresse à l'extérieur du Canada	50,5	46,8	93	1,6	3,5	1,6	3,3	1,4	3,1
Résidence									
Milieu rural / exploitation agricole	1 213,0	1 112,3	92	75,3	6,8	73,1	6,6	70,7	6,4
Centre de population									
Petit (1 000 à 29 999 habitants)	600,3	560,6	93	38,2	6,8	36,8	6,6	35,5	6,3
Moyen (30 000 à 99 999 habitants)	425,6	405,4	95	24,7	6,1	23,5	5,8	22,9	5,7
Grand (100 000 habitants ou plus)	2 707,2	2 574,5	95	121,4	4,7	115,6	4,5	112,3	4,4

% admissibles pour le couplage = (nombre de répondants au recensement [questionnaire détaillé] ayant une clé de couplage unique / nombre de répondants recensés à l'aide du questionnaire détaillé) * 100
% hospitalisés = (nombre de répondants au recensement [questionnaire détaillé] qui ont été hospitalisés / nombre de répondants recensés à l'aide du questionnaire détaillé admissibles) * 100
Sources : Recensement de la population de 2006; fichier de données résultant du couplage recensement–BDCP pour 2006-2007 à 2008-2009.

Discussion

Selon la méthode d'appariement exact déterministe hiérarchique, environ 80 % des clés de couplage identifiées dans les données hospitalières ont été couplées avec le Recensement de 2006. Ce taux est similaire à celui qui ressort d'autres études canadiennes qui ont signalé un taux de couplage de 75 % parmi les enregistrements pour lesquels on s'attendait à un couplage[5,40,41].

L'approche hiérarchique a permis de trouver des couplages qu'un appariement exact déterministe exécuté en une seule itération aurait manqués[13]. Bien que la majorité des couplages réalisés aux fins de la présente étude ont eu lieu au cours de la première itération, un autre 8 % à 17 % se sont produits lors d'itérations subséquentes. L'utilisation de codes postaux mis à jour à partir des données d'impôt afin de tenir compte de la mobilité a entraîné d'autres couplages avec les données hospitalières, particulièrement dans le cas des années les plus récentes, et a permis de franchir un obstacle couramment associé à la plupart des études de suivi selon le recensement[42]. Les taux de couplage, la couverture et la qualité des couplages de données sont demeurés constants pendant la période de trois ans visée par l'étude.

Les taux de couplage avec la BDCP chez la cohorte des personnes dénombrées à l'aide du questionnaire détaillé

ont été de 5 % à 6 %, ce qui représente le pourcentage ayant été hospitalisées au moins une fois au cours des trois années. Les taux de couplage étaient plus élevés chez certains groupes, à savoir les personnes âgées, les personnes appartenant aux quintiles inférieurs de revenu et les Autochtones. Cette observation concorde avec celles de travaux antérieurs[36,38,39], et vient étayer la validité du processus de couplage ainsi que la pertinence d'utiliser des données couplées dans les analyses de la santé.

L'analyse de la couverture a révélé que le fichier de données couplées recensement–BDCP tient compte de la majorité des hospitalisations survenues au cours de la période visée par l'étude (résultats pondérés). De plus, les profils d'hospitalisation selon les caractéristiques des patients étaient similaires aux résultats obtenus à partir de la simple BDCP, à quelques exceptions près.

Les taux de couverture non pondérés variaient selon la région géographique, les territoires, le Manitoba et la Saskatchewan présentant les taux les plus élevés. Cela tient à la stratégie d'échantillonnage utilisée pour le recensement, en fonction de laquelle jusqu'à 100 %

Tableau 4
Taux de couverture non pondérés et pondérés pour les congés obtenus d'un hôpital de soins de courte durée, selon la province, le sexe et le groupe d'âge, Canada (sauf le Québec), 2006-2007, 2007-2008 et 2008-2009

	2006-2007					2007-2008					2008-2009				
	Sorties d'hôpital non couplées (dénominateur)	Répondants au recensement [questionnaire détaillé] hospitalisés (numérateur)				Sorties d'hôpital non couplées (dénominateur)	Répondants au recensement [questionnaire détaillé] hospitalisés (numérateur)				Sorties d'hôpital non couplées (dénominateur)	Répondants au recensement [questionnaire détaillé] hospitalisés (numérateur)			
		Non pondérés		Pondérés			Non pondérés		Pondérés			Non pondérés		Pondérés	
	milliers	milliers	%	milliers	%	milliers	milliers	%	milliers	%	milliers	milliers	%	milliers	%
Canada (sauf le Québec)	2 045,0	351,8	17,2	1 629,4	79,7	1 967,5	338,2	17,2	1 566,0	79,6	1 942,1	327,0	16,8	1 513,4	77,9
Province/Territoire															
Terre-Neuve-et-Labrador	52,9	8,4	16,0	40,5	76,5	51,0	8,4	16,4	39,8	78,0	49,4	8,0	16,3	38,3	77,6
Île-du-Prince-Édouard	15,8	2,5	16,0	12,9	81,6	15,1	2,5	16,5	12,6	83,7	14,8	2,3	15,5	11,7	79,6
Nouvelle-Écosse	80,4	13,3	16,5	65,8	81,8	76,7	12,7	16,6	63,1	82,3	77,0	12,5	16,2	62,4	81,0
Nouveau-Brunswick	87,1	14,2	16,3	69,7	80,0	82,2	13,4	16,3	66,1	80,4	79,3	12,6	15,9	61,8	78,0
Ontario	917,7	149,5	16,3	741,8	80,8	876,3	143,0	16,3	711,3	81,2	863,7	138,6	16,0	688,0	79,7
Manitoba	113,8	25,3	22,2	91,6	80,4	109,6	24,0	21,9	86,6	79,0	109,0	23,6	21,7	84,9	77,9
Saskatchewan	121,6	25,9	21,3	98,8	81,3	115,7	24,5	21,2	92,7	80,1	113,3	23,6	20,8	89,5	79,0
Alberta	291,0	49,7	17,1	227,4	78,2	282,9	48,0	17,0	219,4	77,6	279,3	46,0	16,5	210,9	75,5
Colombie-Britannique	352,6	57,4	16,3	273,1	77,4	346,5	56,1	16,2	266,6	76,9	345,0	54,4	15,8	257,8	74,7
Yukon	3,3	1,0	29,0	2,0	62,4	3,3	1,0	31,4	2,2	67,7	3,4	1,0	30,6	2,4	71,5
Territoires du Nord-Ouest	5,1	2,3	45,8	3,3	65,7	5,1	2,3	45,7	3,3	65,8	4,7	2,1	45,0	3,3	69,5
Nunavut	3,7	2,3	63,0	2,4	65,7	3,2	2,2	67,4	2,2	69,4	3,2	2,2	69,0	2,3	71,6
Sexe															
Hommes	858,5	145,6	17,0	679,1	79,1	816,5	139,9	17,1	652,6	79,9	804,4	136,6	17,0	638,0	79,3
Femmes	1 186,5	206,2	17,4	950,3	80,1	1 151,0	198,3	17,2	913,5	79,4	1 137,7	190,4	16,7	875,4	76,9
Groupe d'âge															
Moins de 1 an	54,3	8,9	16,5	39,1	72,0	0,2	0,0	12,6	0,1	55,4					
1 an à 4 ans	43,9	8,1	18,4	34,2	78,0	34,7	5,9	17,1	25,4	73,2	20,4	3,3	16,2	14,3	70,3
5 à 9 ans	27,4	5,1	18,7	22,7	82,9	26,4	5,0	18,8	21,9	83,0	25,3	4,5	17,7	19,5	77,3
10 à 14 ans	29,7	5,6	18,7	23,9	80,3	29,3	5,5	18,8	23,4	80,0	28,7	5,1	17,7	21,8	75,9
15 à 19 ans	63,7	11,9	18,7	47,4	74,3	63,9	12,0	18,7	47,1	73,6	62,5	11,7	18,7	45,2	72,3
20 à 24 ans	99,4	16,8	16,9	69,8	70,2	99,7	15,7	15,7	63,0	63,2	97,1	14,6	15,0	57,4	59,1
25 à 34 ans	284,1	48,7	17,2	223,3	78,6	288,8	46,8	16,2	214,4	74,2	286,5	43,4	15,2	196,6	68,6
35 à 44 ans	211,8	36,6	17,3	168,7	79,7	207,8	34,9	16,8	159,7	76,9	202,1	32,8	16,3	149,1	73,8
45 à 54 ans	222,1	38,9	17,5	178,8	80,5	222,3	38,9	17,5	179,0	80,5	221,7	38,0	17,1	175,7	79,2
55 à 64 ans	261,5	47,4	18,1	220,2	84,2	264,0	48,1	18,2	222,5	84,3	266,5	47,2	17,7	221,0	82,9
65 à 74 ans	289,0	51,6	17,8	245,2	84,9	285,7	51,9	18,2	246,7	86,3	286,4	51,7	18,0	245,3	85,7
75 ans et plus	458,2	72,1	15,7	356,2	77,7	444,7	73,6	16,5	362,8	81,6	445,1	74,7	16,8	367,5	82,6

Taux de couverture non pondéré = (Total des hospitalisations chez les répondants au recensement [questionnaire détaillé] / Total des hospitalisations selon la BDCP) * 100
Taux de couverture pondéré = (Total des hospitalisations chez les répondants au recensement [questionnaire détaillé] et poids du recensement appliqués / Total des hospitalisations selon la BDCP) * 100

Notes : Les estimations non couplées et couplées de la BDCP ne tiennent pas compte des mortinaissances, des hospitalisations dans d'autres genres d'établissements que de soins de courte durée, des non-Canadiens, des naissances survenues après le jour du recensement (16 mai 2006), des cartes d'assurance-maladie non émises par la province ou le territoire en question, y compris inconnu, CA, 99, ni des patients transférés d'une résidence pour personnes âgées.
Pour les estimations non couplées de la BDCP, on identifie la province à l'aide du code health_card_prov_code.
Sources : Recensement de la population de 2006; fichier de données résultant du couplage recensement–BDCP pour 2006-2007 à 2008-2009.

Ce que l'on sait déjà sur le sujet

- Le couplage d'enregistrements est une façon efficace sur le plan des coûts d'ajouter de la valeur aux ensembles de données administratives et d'enquêtes.
- La combinaison de sources de données complémentaires permet des analyses qui autrement ne seraient pas possibles.
- Le recours aux registres provinciaux d'assurance-maladie aux fins de couplages d'enregistrements pour étudier la santé de la population est une méthode bien établie.
- Les méthodes fondées et non fondées sur des registres pour le couplage des données administratives sur la santé et des données du Recensement de 2006 ont produit des données d'une qualité suffisante pour la recherche, en ce qui a trait aux provinces de l'Ontario et du Manitoba, à l'égard de l'exercice 2006-2007.

Ce qu'apporte l'étude

- Cette étude présente les résultats d'un couplage par appariement exact déterministe hiérarchique de données du Recensement de 2006 et de données tirées d'enregistrements de la Base de données sur les congés des patients (BDCP), pour les exercices 2006-2007 à 2008-2009. Le couplage porte sur tous les secteurs de compétence du Canada (sauf le Québec) et se fonde sur la date de naissance, le code postal et le sexe.
- Le recours aux données sur l'impôt permet d'avoir des codes postaux à jour pour les personnes dénombrées au recensement, ce qui a pour but de réduire au minimum le nombre de couplages qui pourraient être manqués.
- La cohorte ayant résulté du couplage représente environ 80 % des hospitalisations survenues au cours de la période 2006-2007 à 2008-2009.
- Le fichier issu du couplage recensement–BDCP, qui est représentatif de la population nationale, comble des lacunes statistiques et offre de nouvelles possibilités de recherche.

des répondants au recensement vivant dans les régions éloignées et dans les réserves indiennes et établissements indiens devaient remplir le questionnaire détaillé. Une fois la pondération effectuée, les taux de couverture pour le Manitoba et la Saskatchewan étaient plus près de ceux observés pour les autres provinces, mais les taux pondérés étaient tout de même plus faibles pour les territoires que pour les provinces. Il est possible que les taux d'hospitalisation des habitants des territoires soient sous-estimés dans les données couplées en raison du sous-dénombrement, de leurs taux de mobilité élevés et(ou) de la tendance chez eux à être hospitalisés hors de leur territoire de résidence[22,24].

Comme on s'y attendait, l'analyse a révélé un sous-dénombrement potentiel des cas d'hospitalisation selon certains groupes d'âge. Dans le cas des nourrissons âgés de moins de 1 an, les faibles taux de couplage et de couverture pourraient être attribuables à l'absence de NAM dans les enregistrements de naissance des hôpitaux[43]. Malgré que l'on ait tenté au cours du traitement des données de résoudre ces cas, certains n'ont probablement pas été résolus. Les taux de couverture étaient faibles également chez les 15 à 24 ans, ce qui reflète peut-être le sous-dénombrement au recensement des personnes ayant une situation de vie moins stable et(ou) le dénombrement incomplet de certaines populations autochtones[22,44]. Étant donné que ces populations peuvent présenter des taux d'hospitalisation relativement élevés, les couplages manqués ont peut-être une incidence plus importante sur les taux de couverture que ne le suggèrent les nombres absolus[38].

Les jumeaux de même sexe résidant à la même adresse le jour du recensement et les jumeaux hospitalisés de même sexe auxquels correspondait le même code postal ont été supprimés du fichier de données couplées en raison de leurs clés de couplage non uniques. Cela étant dit, les naissances multiples (jumeaux ou plus) représentent environ 3 % de toutes les naissances annuellement, et celles de trois enfants ou plus de même sexe sont plus rares encore[45].

Limites

Le fichier recensement–BDCP couplé comporte plusieurs limites importantes.

Premièrement, pour obtenir un ensemble de clés de couplage uniques, il a été nécessaire de retirer certaines clés des fichiers du recensement et de la BDCP. Dans le cas du questionnaire abrégé, 97,6 % des clés ont été conservées; dans celui du questionnaire détaillé, on en a conservées environ 94 %. Le taux d'admissibilité plus faible chez les personnes ayant rempli le questionnaire détaillé du recensement ressortait le plus chez les personnes ayant un faible statut socioéconomique, les personnes d'identité autochtone, les personnes vivant en milieu rural ou dans une exploitation agricole et les habitants du Nunavut et de la Colombie-Britannique. Une telle situation tient à des facteurs tels que l'enregistrement inexact ou incomplet de la date de naissance et(ou) d'autres renseignements démographiques[5,15,34,46].

Deuxièmement, aux fins d'analyse de la couverture, on a comparé les hospitalisations dans le fichier recensement–BDCP couplé à celles dans les seules données de la BDCP. Les populations visées diffèrent toutefois. Bien que l'on ait tenté de retirer les enregistrements de la BDCP qui correspondaient à une personne non dénombrée lors du recensement (questionnaire détaillé), il est possible que certains y soient demeurés.

Enfin, le Québec est exclu de la présente étude parce que les données hospitalières pour cette province ne sont pas fournies à Statistique Canada. Les hospitalisations au Québec de personnes vivant à l'extérieur du Québec ne sont d'ailleurs pas reflétées dans les données couplées.

Mot de la fin

L'échantillon qui est représentatif de la population nationale et la puissance statistique qu'offre le fichier de données couplées, compte tenu de sa taille et de sa couverture, fournissent de nouvelles occasions de recherche. L'analyse des taux annuels de couplage et de couverture donne à penser que l'intégralité et la

qualité du fichier sont demeurées constantes au fil du temps. Il conviendrait que les chercheurs qui utiliseront ces données couplées tiennent compte de l'incidence éventuelle des méthodes de couplage, des différences d'admissibilité pour le couplage, des taux de couplage et de couverture ainsi que des groupes de populations exclues de l'étude.

Références

1. W.E. Winkler, Record linkage, *Sample Surveys: Design, Methods and Applications*, 29A, 2009, p. 351-380.
2. L. Lix, M. Yogendran, M. Burchill *et al.*, *Defining and Validating Chronic Diseases: An Administrative Data Approach*, Winnipeg, Manitoba, Manitoba Centre for Health Policy, 2006.
3. K. Iron, B.M. Zagorski, K. Sykora *et al.*, « Living and dying in Ontario: An opportunity for improved health information », *ICES Investigative Report*, Toronto, Institute for Clinical Evaluative Sciences, 2008.
4. V.W. Leong, S. Morgan, S.T. Wong *et al.*, « Registration for public drug benefits across areas of differing ethnic composition in British Columbia, Canada », *BMC Health Services Research* (électronique 10(171), 2010, disponible à l'adresse http://www.biomedcentral.com/1472-6963/10/171
5. R. Wilkins, M. Tjepkema, C. Mustard *et al.*, « Études canadienne de suivi de la mortalité selon le recensement, 1991 à 2001 inclusivement », *Rapports sur la santé*, 19(3), 2008, p. 27-48.
6. K. Wilkins, M. Shields et M. Rotermann, « Utilisation des hôpitaux de soins de courte durée par les fumeurs – Une étude prospective », *Rapports sur la santé*, 20(4), 2009, p. 83-93.
7. C.D. Holman, A.J. Bass, D.L. Rosman *et al.*, « A decade of data linkage in Western Australia: Strategic design, applications and benefits of the WA data linkage system », *Australian Health Review*, 32(4), 2008, p. 766-767.
8. E. Ng, « Effet de l'immigrant en santé et taux de mortalité », *Rapports sur la santé*, 22(4), 2011, p. 25-30.
9. S.W. Hwang, R. Wilkins, M. Tjepkema *et al.*, « Mortality among residents of shelters, rooming houses, and hotels in Canada: 11-year follow-up study », *British Medical Journal*, 339, 2009, p. b4036. doi:101136/bmj.b4036
10. M. Tjepkema et R. Wilkins, « Espérance de vie restante à l'âge de 25 ans et la probabilité de survie jusqu'à l'âge de 75 ans, selon la situation socioéconomique et l'ascendance autochtone », *Rapports sur la santé*, 22(4), 2011, p. 31-36.
11. D.G. Manuel, R. Perez, C. Bennett *et al.*, « Sept années en plus : Incidence du tabagisme, de la consommation d'alcool, de l'alimentation, de l'activité physique et du stress sur la santé et l'espérance de vie en Ontario », un rapport de Santé publique Ontario et de l'Institute for Clinical Evaluative Sciences, 2012.
12. R. Fransoo, P. Martens, H. Prior *et al.*, *Adult Obesity in Manitoba: Prevalence, Associations, and Outcomes*, Winnipeg, Manitoba, Manitoba Centre for Health Policy, 2011.
13. J. Bernier et L. Nobrega, « Un aperçu du couplage d'enregistrements », *Recueil – Symposium 99 : combiner des données de sources différentes, 1999*, Ottawa, Ministre de l'Industrie, 2000.
14. R. Karmel, P. Anderson, D. Gibson *et al.*, « Empirical aspects of records linkage across multiple data sets using statistical linkage keys: the experience of PIAC cohort study », *BMC Health Services Research*, 10, 2010, p. 41, disponible à l'adresse http://www.biomedcentral.com/1472-6963/10/41
15. NCSIMG, *Statistical Data Linkage in Community Services Data Collection*, Canberra, Australia, Australian Institute of Health and Welfare, 2004, disponible à l'adresse http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=10737419651&libID=10737419650
16. B. Li, H. Quan, A. Fong *et al.*, « Assessing record linkage between health care and Vital Statistics databases using deterministic methods », *BMC Health Services Research* (électronique) 2006, 6(48), disponible à l'adresse http://www.biomedcentral.com/1472-6963/6/48
17. Australian Institute of Health and Welfare, *Comparing an SLK-based and a Name-based Data Linkage Strategy: An Investigation into the PIAC Linkage*, Data Linkage Series No. 11 (Catalogue CSI 11) Canberra Australia, Australian Institute of Health and Welfare, 2011.
18. HE Development Team, *Methodology for creation of the HES patient ID (HESID)*, disponible à l'adresse http://www.hscic.gov.uk/media/1370/HES-Hospital-Episode-Statistics-Replacement-of-the-HES-patient-ID/pdf/HESID_Methodology.pdf
19. M. Rotermann, C. Sanmartin, R. Trudeau *et al.*, « Couplage de données de recensement et de données d'hôpital : deux approches », Rapports sur la santé, 25(10), 2014, p. 3-15.
20. Statistique Canada, Couplages d'enregistrements approuvés, disponible à l'adresse http://www.statcan.gc.ca/fra/enregistrement/somm
21. Statistique Canada, *Directive sur le couplage d'enregistrements*, disponible à l'adresse http://www.statcan.gc.ca/fra/enregistrement/politique4-1
22. Statistique Canada, *Rapport technique du Recensement de 2006 : Couverture*, Ottawa, Ministre de l'Industrie, 2010. disponible à l'adresse http://www12.statcan.gc.ca/census-recensement/2006/ref/rp-guides/rp/coverage-couverture/cov-couv_index-fra.cfm
23. Statistique Canada, *Recensement de la population de 2006*, disponible à l'adresse http://www12.statcan.gc.ca/census-recensement/2006/index-fra.cfm
24. Statistique Canada, *Rapport technique du Recensement de 2006 : Échantillonnage et pondération*, Ottawa, Ministre de l'Industrie, 2009, disponible à l'adresse http://www12.statcan.gc.ca/census-recensement/2006/ref/rp-guides/rp/sw-ep/pdf/92-568-X2006001-fra.pdf
25. G. Roberts, « Analyser les microdonnées du recensement dans un CDR : quelle pondération utiliser? » *Le Bulletin technique et d'information des Centres de données de recherche* (Statistique Canada, nº 12-002-X au catalogue) automne 2012, 5(1), disponible à l'adresse http://www.statcan.gc.ca/pub/12-002-x/2012001/article/11642-fra.htm
26. Institut canadien d'information sur la santé, *Métadonnées de la Base de données sur les congés des patients (BDCP)*, disponible à l'adresse http://www.cihi.ca/cihi-ext-portal/internet/fr/document/types+of+care/hospital+care/acute+care/dad_metadata
27. Institut canadien d'information sur la santé, *Étude de l'ICIS sur la qualité des données de la Base de données sur les congés BDVP des patient pour 2006-2007, 2007-2008, et 2008-2009*, disponible à l'adresse https://secure.cihi.ca/estore/productSeries.htm?locale=fr&pc=PCC228

28. St- Pierre M, Gagné W. *Création d'un prototype de la base de la population canadienne*. Statistics Canada internal working documentation. Ottawa: Statistics Canada, 2012.

29. Institut canadien d'information sur la santé, *Indicateurs de santé 2013*. Ottawa, Institut canadien d'information sur la santé, 2013, disponible à l'adresse https://secure.cihi.ca/free_products/HI2013_Jan30_FR.pdf

30. J.L. Hoving, A. Del Monaco et E. MacFarlance, « Methodological issues in linking study participants to Australian cancer registries using different methods: lessons from a cohort study », *Australian and New Zealand Journal of Public Health*, 29(4), 2005, p. 378-382.

31. Australian Bureau of Statistics, *Assessing the Quality of Linking School Enrolment Records to 2011 Census Data: Deterministic Linkage Methods*, Research Paper 1351.0.55.045, disponible à l'adresse http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/FAD1F2FC23A36CE4CA257C3F00180F64/$File/1351055045_dec%202013.pdf

32. J. Bass et C. Garfield, « Statistical linkage keys: How effective are they? », *Proceedings of the Symposium on Health Data Linkage, 2003*, disponible à l'adresse http://www.publichealth.gov.au/pdf/reports_papers/symposium_procdngs_2003/bass.pdf

33. A. Beauchamp, A.M. Tonkin, V. Sundararajan *et al.*, « Validation of de-identified record linkage to ascertain hospital admissions in a cohort study », *BMC Medical Research Methodology*, 11(42), 2011, doi:10.1186/1471-2288-11-42

34. L.S. Jebamani, C.A. Burchill et P.J. Martens, « Using data linkage to identify First Nations Manitobans », *Canadian Journal of Public Health*, 96(1), 2005, p. S28-32.

35. M. Rotermann « Évaluation de la couverture des données de l'Enquête sur la santé dans les collectivités canadiennes couplées aux dosiers de malades hospitalisés », *Rapports sur la santé*, 20(1), 2009, p. 49-57.

36. Institut canadien d'information sur la santé, *Les soins de santé au Canada*, Ottawa, Institut canadien d'information sur la santé, 2011. disponible à l'adresse http://secure.cihi.ca/freeproducts/HCIC_2011_seniors_report_FR.pdf

37. H. Trottier, L. Martel, C. Houle *et al.*, « Vieillir chez soi ou en établissement : à quoi cela tient-il? », *Rapports sur la santé*, 11(4),2000, p. 55-68.

38. G. Carrière, R. Garner, C. Sanmartin *et al.*, « Hospitalisations dans des hôpitaux de soins de courte durée et identité autochtone au Canada, 2001-2002 », *Série de documents de travail de la recherché sur la santé* (nº 82-622-X au catalogue) Ottawa, Statistique Canada, 2010, disponible à l'adresse http://www.statcan.gc.ca/pub/82-622-x/82-622-x2010005-fra.pdf

39. C. Sanmartin, S. Khan et l'équipe de recherche de l'Initiative sur les données longitudinales administratives et sur la santé, « Hospitalisations pour des conditions propices aux soins ambulatoires (CPSA) : les facteurs qui importent », *Série de documents de travail de la recherché sur la santé* (nº 82-622-X au catalogue) Ottawa, Statistique Canada, 2011. disponible à l'adresse http://www.statcan.gc.ca/pub/82-622-x/82-622-x2011007-fra.pdf

40. J. Brennan et M. Lessard, *Une exemple de couplage d'enregistrements: Couplage du fichier des immigrants reçus aux données du TIFF*, Séminaire de la DMEM de Statistique Canada, le 9 avril, 2013.

41. C. Houle, J.M. Berthelot, P. David P *et al.*, « Le projet d'appariement du Recensement et des fichiers de soins de santé du Manitoba », *Federal Committee on Statistical Methodology Papers* 1997, disponible à l'adresse http://www.statcan.gc.ca/pub/11f0019m/11f0019m1996091-fra.pdf

42. T. Blakely, A. Woodward et C. Salmond, « Anonymous linkage of New Zealand mortality and census data », *Australian and New Zealand Journal of Public Health*, 24(1), 2000, p. 92-95.

43. Institut canadien d'information sur la santé, tableaux spéciaux : *Availability of newborn health card numbers 2006/07 to 2012/13*, reçu le 16 avril 2014.

44. Statistique Canada, *Réserves indiennes et établissements indiens partiellement dénombrés*, disponible à l'adresse http://www12.statcan.gc.ca/census-recensement/2006/ref/notes/aboriginal-autochtones-fra.cfm

45. Statistique Canada, Naissances 2009, disponible à l'adresse http://www.statcan.gc.ca/pub/84f0210x/84f0210x2009000-fra.pdf

46. Australian Institute of Health and Welfare, Australian Bureau of Statistics, *National Best Practice Guidelines for Data Linkage Activities Relating to Aboriginal and Torres Strait Islander People*, Canberra, Australia, Australian Institute of Health and Welfare, 2012, disponible à l'adresse http://www.aihw.gov.au/publication-detail/?id=60129542148